**Metabolomics Approach to Identify Molecules and Pathways Involved in the Development of Atherosclerotic Coronary Artery Disease – a RTI RCMRC Pilot Study**

Metabolomic Analysis: NIH Eastern Regional Comprehensive Metabolomics Resource Core (RTI RCMRC)

PI, RTI RCMRC Core Collaboration: Xuming Dai MD, PhD

IRB Number(s):

**Abstract:**

Genetics play major roles in the development of atherosclerotic coronary artery disease (CAD). Despite tremendous efforts worldwide invested to decipher the genetic components controlling the development of CAD, the genetic architecture of CAD remains largely unclear. As part of an on-going effort to identify molecules and pathways involved in the development of atherosclerotic CAD, we propose to use rigorous angiographic criteria to define CAD phenotype for genomics and metabolomics study. We identified two extreme groups, namely “young CAD” group, who are very young individuals (age <= 40 years) proven to have severe CAD required revascularization, and “CAD-free elderly”, who are at very advanced age (Age >= 80 years) but have no angiographically apparent CAD. Phenotypically, these two groups are in sharp contrary. Conventional risk factors account for small portion of different phenotypes. We hypothesize that there are genetically programmed pathways and molecules accelerating atherosclerotic pathogenesis, in the “young CAD” patients and preventing the development of CAD in the “CAD-free elderly” patients. We sought to combine genomics and metabolomics approaches to profile and identify these pathways and molecules. Both plasma and urine samples from patients in these two groups, and their age matched control groups, will undergo unbiased metabolomics profiling with high throughput quantitative nuclear magnetic resonance (NMR) and mass spectrometry (MS) technology in RTI metabolomics core facility. Comprehensive statistic and multi-variant analytic approaches will be used to identify pathways and molecules significance to the pathogenesis of atherosclerosis. These data will be integrated with genomics data from next generation sequencing of genetic materials from the same groups of patients to further explore the molecular mechanisms underlying atherosclerosis and CAD.

**Sample Description:**

Aliquots of each de-identified plasma sample were shipped to the NIH RTI-RCMRC on dry ice and immediately stored at -80 °C after being logged in for metabolomics analysis. A total of 106 study samples were thawed on ice for sample preparation. 350 uL of plasma sample were thawed and transferred to labeled tubes on ice where they were mixed with 350 uL Saline master mix (2mM Formate). Analytical quality control (QC) phenotypic pooled samples were generated by transferring a 15 µL of each sample of each respective phenotypical experimental sample into different 2.0 mL tubes. Whole study (total) pools were generated by transferring 350 uL of plasma from each Pool sample into a 2.0 mL tube. The tubes were vortexed for 4 min on a multi-tube vortexer and centrifuged at 16,000 rcf for 4 min. A 600 µl aliquot of the supernatant was transferred into pre-labeled 5mm NMR tubes for data acquisition on a 700 MHz spectrometer.

The data obtained for the NMR metabolomics analysis can be found in the accompanying files:

Procedures: 1. CAD Metabolomics Procedures.docx

Study Design Tables: 2. CAD Metabolomics Study Design Table.xls

Metadata: 3. CAD Metabolomics METADATA.xlsm

Processed Data: 4. CAD Metabolomics Normalized Binned Data.xlsx

Raw Data: 5. CAD Metabolomics NMR Raw Data.zip

**Notes:**

Full sample preparation and analysis procedures are available in the accompanying document entitled **1. CAD Metabolomics NMR Procedures**.

Descriptions of abbreviations for factors are available in the Variable Dictionary in the accompanying file no. **2. CAD Metabolomics NMR Study Design Table.xls**.

The phenotypic and normalized data are available in the accompanying files: **4. CAD Metabolomics NMR Normalized Binned Data.xlsx** for normalized binned NMR data. Sample ID and factors can be found in the first 5 columns and other columns in the spreadsheet contain sample metadata and the normalized binned data. If the statistical program does not allow variable names to begin with a number then add a prefix to the column names, for example, bin\_8.98 instead of 8.98.

The Sample ID serves as the unique identifier (Graphical ID) of the individual samples and is used as the NMR folder name in the raw NMR data file **5. CAD Metabolomics NMR Raw Data.zip**.